

Human Interface

ヒューマンインタフェース学会誌

1999
Vol.1
No.2

■巻頭言

ヒューマンインタフェースとデザイン

■企画記事

インタフェースの心理と生理 (第1回)

コミュニケーションの身体性

インタフェースの鳥瞰図

研究所紹介・「ATR知能映像通信研究所」

■報告

第1回研究会報告

IFAC-MMS'98 ワークショップ報告

国際会議報告

「第14回Technology and Persons with Disabilities」

■案内

ヒューマンインタフェースシンポジウム'99

研究会プログラム

研究会／論文特集号／協賛行事

会誌投稿

入会案内、他

■編集後記

■論文

・横1列に逐次提示された文字と文字音声に時間差がある場合の提示時間最適配分

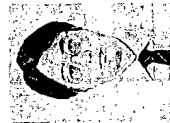
・プラント運転員の異常診断行動のモデリングと
ヒューマンモデルシミュレーションによる人間認知信頼性
(HCR)曲線の導出法の研究

・動的顔画像からのリアルタイム表情認識システムの試作

1-2

著者紹介

Wu Wei (学生会員)



1997年京都大学大学院工学研究科修士課程修了。現在、同大学院博士課程在学中。マンマシンインタフェースにおける人間と機械の相互作用に関する研究に興味を持つ。日本原子力学会、ヒューマンインタフェース学会の学生会員。

中川 隆志 (正会員)



1991年京都大学第三学群基礎工学類卒業。1993年同大学院工学研究科修士課程修了。同年三菱電機(株)に入社、産業システム研究所に勤務。主に発電プラントのヒューマンファクター、ヒューマンインタフェース技術の研究に従事。1997年人工知能学会研究奨励賞、1997年電気学会論文発表賞受賞。ヒューマンインタフェース学会、計測自動制御学会、日本原子力学会の会員。

吉川 榮和 (正会員)



1970年京都大学大学院工学研究科電気工学第二学専攻修士課程修了。同年同大学原子エネルギー研究所助手。1974年動力炉・核燃料開発事業団助役。1981年京都大学原子エネルギー研究所助教授。1992年同教授。1996年同大学大学院エネルギー科学研究科教授。現在に至る。エネルギー情報学、マンマシンシステムの研究に従事。1992年システム制御情報学会植木記念賞論文賞、1995年計測自動制御学会論文賞受賞。電気学会、計測自動制御学会、日本原子力学会などの会員。工学博士。

原著論文

動的顔画像からのリアルタイム表情認識システムの試作

下田 宏*1 中川 隆志 吉川 榮和*1

A Prototype of a Real-time Expression Recognition System from Dynamic Facial Image
Hiroshi Shimoda,*1 Takeshi Kunihira*1 and Hidekazu Yoshikawa*1

Abstract — A human-machine interface(HMI) using facial expression is expected to realize higher level cooperation between human and machine, because this idea will develop the HMI towards more natural human-machine communication. In this paper, we first introduce an automatic and real-time recognition method of facial expression, and then describe a prototype system developed based on this method. The system captures dynamic facial images from a color CCD camera, extracts feature points on facial parts (eyebrows, eyes and mouth) by real-time image processing, calculates feature values corresponding to the AU of FACS, and recognizes the strength of neutral and six basic expressions for each part of the face (upper, middle and lower) by employing fuzzy estimation.

Keywords : facial expression recognition, FACS, image processing, fuzzy estimation

1. はじめに

コンピュータに代表される機械が人間の生活に浸透し、人間と機械の接する機会が増加するにつれて、どのようなシステムにすればより快適に機械と接することができるかが研究課題とされるようになってきた。現在のところ、人間と機械とがコミュニケーションを行う場合には、歩み寄るのとはほとんどが人間の側であり、人間にとつて快適であるとしても言えない状況である。人間と機械とのインタフェース、すなわちヒューマンマシンインタフェースをより自然なものにするためには、従来のように人間の側から機械に働きかけるだけでなく、機械の方でも人間の状態を察知し、それに即した対応ができるような、適応型のインタフェースを構築する必要がある。

人間と人間とが向かい合ってコミュニケーションを行う場合、言語だけでなく顔の表情、ジェスチャ、声の抑揚などたくさんの方で情報を伝達される。なかでも、顔の表情は非常に重要な役割を担っており、人間は相手の表情を読み取りそこから心理状態を推定することにより、円滑なコミュニケーションを行っている。したがって、人間と機械との交流の場合においても、その円滑なコミュニケーションを実現するためには、機械が人間から受け取る情報の一つとして、表情は欠かせないものであると考えられる。ヒューマンマシンインタフェースに表情を利用し、機械が人間の顔面状態を推定することにより、より高次の協調関係を表現できることが期待される。

*1: 京都大学大学院エネルギー科学研究科

*1: Graduate School of Energy Science, Kyoto University

人間同士のコミュニケーションにおいては、人間は相手の表情を動的にとらえており、時間とともに微妙に変化していく表情の流れから相手の顔面状態を読み取っている。そのため、動的な顔画像から表情やその変化のイメージング、選定を導くことが重要であると考えられる。そこで本研究では、人間と機械とのよりよいコミュニケーション方法の実現を目指し、その基礎研究として動的な顔画像からリアルタイムで表情を認識するシステムを試作することを目的とする。特に、ヒューマンマシンインタフェースに表情認識を応用するためには、その認識処理のリアルタイム化が重要であると考え、認識精度に加えて認識速度の向上に主眼をおいて研究を進めた。従来の機械による自動表情認識のアプローチとしては、(1)顔全体を一つのパターンとしてとらえる方法[1]、(2)顔画像のオプティカルフローから目や口などの各器官の変化をとらえる方法[2],[3]、(3)SNAKEと呼ばれる変形ベクトルを用いて各器官を抽出しその変化をとらえる方法[4],[5]、の三つに大別できる。(1)の代表的なものとしては、ポテンシャルネットに基づくものがあるが、顔の各器官を個別に解析しなくともよい反面、顔の個人差に左右されるという問題点がある。一方、(2)や(3)の方法では、個人差に依存しないという特徴があるものの、処理時間が長くなるため、リアルタイムでの表情認識には適していなかった。本研究では、従来研究の問題点を克服し、主に認識処理時間の短縮を目指して、カラーCCDカメラで撮像された顔の動画画像から色情報と濃淡情報を用いて顔輪郭や各器官を抽出し、そこからリアルタイムで表情を認識する手法を提案する。

以下では、心理学分野における表情の研究成果をふま

えて、顔画像から表情を認識する手法について説明し、さらに、その手法に基づいて試作した動画面像によるリアルタイム表情認識システムについて述べる。

2. 心理学分野の表情研究

心理学の分野では、カウセンリングなどへの応用を目的として、古くから表情に関する様々な研究が行われてきた。本研究ではその成果を積極的に活用した表情推定手法を提案するが、主に表情および感情に関連する以下の3つの研究成果を利用している。

- Facial Action Coding System (FACS)
- 顔の3領域
- 基本6表情

本章では、これらについて簡単にまとめる。

2.1 FACS

Facial Action Coding System (FACS) は、Ekman と Friesen によって提案された体系的な表情記述法であり、表情に関連する研究において最も広く利用されている。FACS では、表情を Action Unit (AU) と呼ばれる解剖学的に独立し、かつ視覚的に識別可能な表情動作の最小単位によって記述する。AU は合計40ありあり、これらの組み合わせにより人間のあらゆる表情が記述可能であると言われている。表1にAUの例を示す。たとえば、頬を持ち上げる動作を表す AU6 と唇の端を引っ張り上げる動作を表す AU12 を組み合わせると「微笑み」の表情を表現することができる。後述の基本6表情と AU との関連は、Ekman と Friesen によって明らかにされており、また、基本的に、文化、性別、年齢、個人差に影響されない^[6]。ただし、FACS は本来、人間が表情を認識し分析することを目的として提案されたものであり、AU には定性的な表現が用いられている。

表1 AU (Action Unit) の例
Table 1. Examples of AU (Action Unit)

AU 番号	動作
1	眉の内側を持ち上げる
2	眉の外側を持ち上げる
4	眉を内側に下げる
5	上唇を上げる
6	頬を持ち上げる
12	唇端を引っ張り上げる

2.2 表情筋と顔の3領域

表情変化は、表情筋の収縮に基づくものである。表情筋は主なものだけでも20種類以上あり、それらの収縮・弛緩が複雑に組み合わさって表情を形成している。なかでも重要なものは、前頭筋、眼輪筋、口輪筋、頬筋、頬筋、口角筋などである^[2]。顔は表情筋の影響範囲・移動

るかを推定する。表情の種類としては基本6表情に中立の表情を加えた計7種類を用いる。具体的には、図1節で例を挙げたような各表情に特徴的な表情動作をモデルベースで識別する方法だけでなく、ニューラル・ネットワーク^[9]やパターン・マッチングを利用する方法などが考えられる。後述の試作システムでは、ルールに基づくファジィ推論により各表情に特徴的な表情動作を認識する手法を用いている。

4. 表情認識システムの試作

本章では、前章で述べた手法をもとに試作した、動画面像からリアルタイムで表情を認識するシステムについて述べる。

4.1 システム構成

システム構成の概略を図1に示す。試作したシステムのハードウェアは Graphics Workstation (Silicon Graphics 製、O2、以下 WS)、CCD カメラ (Sony 製、CCD-MCI)、およびディスプレイ (Silicon Graphics 製) によって構成されている。WS に内蔵されているビデオ処理システムを介して、CCD カメラで撮影した顔画像を WS 上に取り込み、特徴量抽出、表情認識の処理を行い、その処理過程、および結果をディスプレイに表示する。また、本システムのソフトウェアは、本研究で作成した (1) 特徴量抽出モジュール、(2) 表情認識モジュール、WS に付属の (3) SGI Video Library (VL)、および (4) グラフィックス表示ライブラリ OpenGL によって構成されている。CCD カメラから得られる画像は、(3) VL を用いて R、G、B、各8ビットの RGB データとして WS のメモリに格納される。その RGB 画像から (1) 特徴量抽出モジュールで特徴量の抽出を行い、得られた特徴量から (2) 表情認識モジュールで表情の認識を行う。また、その過程、および結果を (4) OpenGL を用いてディスプレイに表示する。以下、4.2 節ではシステムの入力となる動画面像について、4.3 節では特徴量抽出モジュールについて、4.4 節では表情認識モジュールについて詳述する。

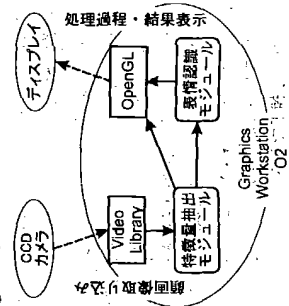


図1 システム構成
Fig. 1 System Configuration

4.2 入力画像

本試作システムの入力となる画像は、1秒間に30フレームの撮影が可能な CCD カメラで撮影されたカラーの動画面像であり、WS に取り込まれた時点での解像度は、320 × 240 ピクセルである。本研究の表情認識手法は、例えば、計算機を操作しているユーザーの顔表情を認識するような場合を想定している。そのため、CCD カメラをユーザーの正面に取り付けることにより、撮像される画像中には、表情認識の対象となる顔がある程度の大きさでほぼ正面向きで写っているものと仮定している。

4.3 特徴量抽出モジュール

試作したシステムの特徴量抽出モジュールは、顔器官上に設定した特徴点を抽出し、特徴点間の距離などから特徴量を算出する方法をとっており、次の4つの部分からなる。

1. 顔領域抽出部
2. 顔器官抽出部
3. 特徴点抽出部
4. 特徴量算出部

ここでは、3.1 節で述べたように、(1) AU をよく表すこと、(2) 画像処理により容易に抽出できること、の2つの条件を考慮して、図2に示すような16個の特徴量を抽出することにした。特徴量を算出するための特徴点としては右眉・左眉について各3点、右目・左目・口について各4点の計16点を選んだ。特徴量抽出モジュールで行う処理の概略を図3に示し、以下で、各部の処理について詳しく述べる。

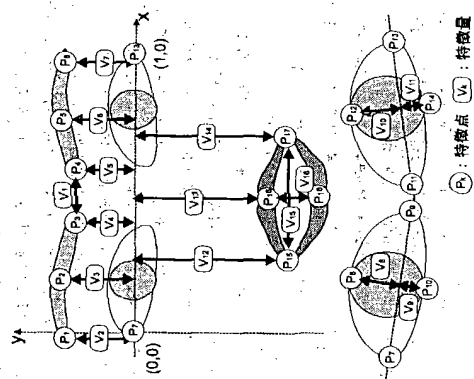


図2 抽出する特徴点と特徴量
Fig. 2 Feature Points and Values to be extracted

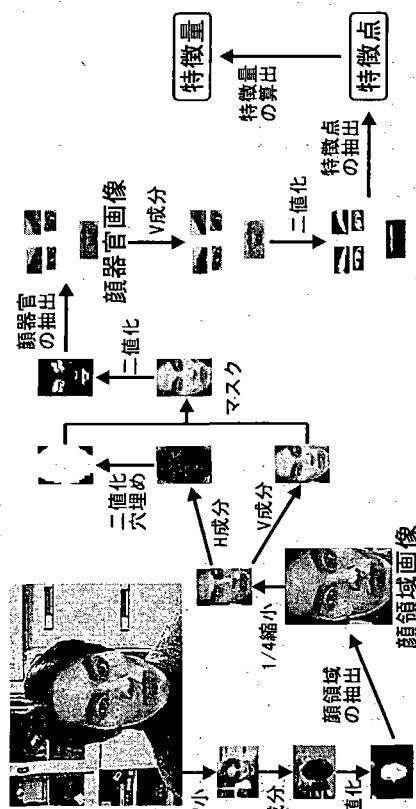


図3 特徴量抽出処理の概略
Fig. 3 Outline of Feature Value Extraction

4.3.1 顔領域抽出部

顔領域抽出部では、入力画像から顔のある位置とその大きさを検出し、顔を含む矩形領域を入力画像から切り出す。具体的には、色情報を用いて肌色の領域を検出し、縦横の比が明らかに顔と異なるものを除き最大の面積を占める領域を顔とみなして抽出する。肌色領域の検出にはHSV表色系(H:hue 色相, S:saturation 彩度, V:value 明度)のH成分を利用している。HSV表色系は、明度の変化による変動が少ないことが実験的に確かめられており、照明環境の変化を受けにくいという利点がある[9]。しかしながら、実際には、H成分だけを手がかりにして顔領域を抽出すると、照明の変化や顔以外の背景の変化により、顔以外の部分が抽出されてしまうことがある。そこで、入力画像が動画であるというところを利用して、前フレームで抽出した顔領域の位置と大きさを参考に、今フレームで抽出した顔領域の位置と大きさを比較し、抽出精度を向上させている。具体的には、入力画像のH成分により画像を二値化し、前フレームで抽出した顔領域の重心から最も近く、かつ、大きさが同程度である領域を抽出している。なお、これらの処理は入力画像を縦横ともに1/8に縮小した画像に対して行っており、これにより処理の大幅な高速化とともに、顔器官やその影の影響を低減してロバスト性を向上させている。

4.3.2 顔器官抽出部

顔器官抽出部では、顔領域の画像から右眉、左眉、右目、左目、口の5つの顔器官を検出し、各顔器官を含む矩形領域を入力画像から切り出す。その際、HSV表色系で明度を表すV成分を利用し、顔器官のつくる暗色部分をとり出すことにより抽出を行う。具体的には、まず顔領域を抽出する場合と同様に、H成分を計算し肌色領域を取り出すことにより、背景を除去するための顔マ

スクをつくる。次に、この顔マスクをV成分を計算した画像に被せてから二値化すると、顔面上にある暗色領域が得られる。これらの領域のうち、形や大きさが不適当なものや除去した後、各領域の位置、幅、高さなどからどの顔器官に対応するかを推定し、その領域を各顔器官とみなして抽出する。なお、ここでは高速化とロバスト性の向上のために顔領域の画像を縦横ともに1/4に縮小した画像に対して処理を行う。

4.3.3 特徴点抽出部

特徴点抽出部では、各顔器官の画像から図2の $P_2\{x|z = 1, \dots, 18\}$ に示した18個の特徴点を抽出する。まず、顔器官の抽出と同様に、HSV表色系のV成分を計算して暗色領域を取り出す。ただし、ここでは高い精度が要求されるため、顔領域や顔器官の抽出の際に行ったような画像の縮小は行わない。暗色領域のうち、各顔器官に対応する領域が最大の面積を占めると考えられるので、それ以外の領域を取り除く。そして、残った領域の端点を検出しそれを基準に特徴点を定める。この時点で、特徴点は入力画像の左上を原点としピクセル単位とするXY座標で表されている。

4.3.4 特徴量算出部

特徴量算出部では、特徴点抽出部で得られる特徴点の座標から特徴量を算出する。ここでは、まず、表情による位置の変化がほとんどない両眼外側の端点である特徴点 P_1 および P_{13} を基準として、各特徴点の座標を図2のようないずれの座標系に変換する。4.2節で述べたように、顔はほぼ正面を向いていることを仮定しており、これにより顔の大きさと傾きを補正できる。

変換した座標系上での各特徴点の座標から次のようにして図2の16個の特徴量を算出する。

表2 特徴量変化と表情の関係

Table 2 Relationship between Feature Value Variation and Expression

表情	上部 ($V_1 \sim V_7$)	中央部 ($V_8, V_9, V_{10} \sim V_{11}$)	下部 ($V_{12} \sim V_{16}$)
驚き	$V_1 \sim V_7$ 増加, V_1 変化無	$V_8, V_9, V_{10} \sim V_{11}$ 増加	V_{12} 増加, $V_{13} \sim V_{16}$ 変化無 or V_{16} 増加, $V_{13} \sim V_{14}, V_{15}$ 変化無
恐れ	V_2, V_4, V_6, V_7 増加, V_3, V_5 減少, V_1 変化無	$V_8, V_9, V_{10} \sim V_{11}$ 増加	V_{12} 増加, $V_{13} \sim V_{16}$ 変化無 or V_{16} 増加, $V_{13} \sim V_{14}, V_{15}$ 変化無
幸福	なし	なし	V_{12}, V_{13} 減少, $V_{14} \sim V_{16}$ 変化無
悲しみ	V_2, V_4 増加, V_1, V_3, V_6 減少, V_5, V_7 変化無	V_8, V_9 増加	V_{12}, V_{14} 増加, V_{13}, V_{15}, V_{16} 変化無
嫌悪	V_1, V_4, V_6 減少, V_2, V_3, V_5, V_7 変化無	V_8, V_9 減少, $V_{10} \sim V_{11}$ 変化無	V_{12}, V_{14} 増加, V_{13} 減少, V_{15}, V_{16} or V_{16} 増加, V_{13} 減少, V_{12}, V_{14}, V_{15} 変化無
怒り	$V_2 \sim V_4$ 減少, V_1 変化無	V_8, V_9 増加, V_{10}, V_{11} 減少, V_{10}, V_{11} 変化無	$V_{12} \sim V_{16}$ 変化無
中立	$V_1 \sim V_7$ 変化無	$V_8, V_9, V_{10} \sim V_{11}$ 変化無	$V_{12} \sim V_{16}$ 変化無

- V_1, V_{15} : x座標の差
- $V_2 \sim V_7, V_{12} \sim V_{14}$: y座標の絶対値
- V_{16} : y座標の差
- V_8, V_9 : 直線 P_1P_2 との距離
- V_{10}, V_{11} : 直線 $P_{11}P_{13}$ との距離

4.4 表情認識モジュール

表情認識モジュールでは、特徴量抽出モジュールで得られた特徴量を顔の3領域のそれぞれについて基本6表情の認識を行う。試作したシステムでは、フuzzy推論を用いて特徴量からの表情認識を行っている。つまり、各表情に特徴的な表情動作を特徴量によって表現し、それをもとにフuzzyルールを作成してフuzzy推論器を構成し、フuzzy推論を行う。各フuzzyルールでは、ある特徴量が中立の表情と比較して大きい(L)か、小さい(S)か、あるいは同じくらい(M)であるかという条件のときに、各表情が強い(ST)か、あるいは弱い(WK)かが記述されている。例えば、表情 exp_1 では、特徴量 V_1 が大きく、 V_2 が小さくなるという場合には、それに対するフuzzyルールは、

IF V_1 is L and V_2 is S

THEN exp_1 is ST and exp_2 are WK

とした。なお、各フuzzy集合のメンバーシップ関数は図4に示すものを用いた。

フuzzy推論器への入力には特徴量を規格化したものを用いる。具体的には、特徴量 V_k のとり得る最大値および最小値をそれぞれ $V_{k,max}$ および $V_{k,min}$ 、中立の表

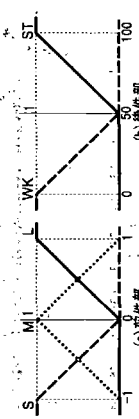


図4 表情認識に用いたメンバーシップ関数
Fig. 4 Membership Function for Expression Recognition



図5 特徴量の範囲測定用画像例
Fig. 5 Example Images for Measurement of Feature Value Range

情での特徴量を $V_{k,neutral}$ とすると、フuzzy推論器への入力 in_k を次式により求める。

$$in_k = \begin{cases} \frac{V_k - V_{k,neutral}}{V_{k,max} - V_{k,neutral}} & (V_k \geq V_{k,neutral} \text{ のとき}) \\ \frac{V_k - V_{k,neutral}}{V_{k,neutral} - V_{k,min}} & (V_k < V_{k,neutral} \text{ のとき}) \end{cases}$$

このときの、 $V_{k,max}$ 、 $V_{k,min}$ および $V_{k,neutral}$ の各値は、図5に示すような、中立の表情の画像や特徴量が最大あるいは最小となっている画像から、特徴量抽出モジュールにより実際に求めた特徴量の値を用いる。

フuzzy推論器からの出力として、各表情の強度が [0, 100] の範囲で得られ、値が大きいくほど強い表情である。このようなフuzzy推論器を、顔の3領域のそれぞれについて作成した。具体的には、2.1節で述べた表情とAUとの関連に基づいて、本研究で提案した16の特徴量の変化と表情の関係を表2のようにルール化した。例えば、驚きの表情の際、顔の上部では、AU1(眉の内側を持ち上げる)、AU2(眉の外側を持ち上げる)の動作があるため、特徴量 $V_2 \sim V_7$ の増加と V_1 の無変化を驚きの表情のルールとした。

4.5 評価実験

ここでは、表情を表出している動画画像から、試作したシステムを用いて表情の認識を行い、その認識精度を評価する実験と結果について述べる。実験は、図6のような照明環境下で、被験者の顔の正面から約80cm離れた

位置に CCD カメラを設置し、試作システムにより被験者の表情を認識した。撮像レートは毎秒 30 フレームである。被験者に奏出してもらった表情は、「中立→幸福→中立→驚き→中立→幸福→中立→驚き」の順に中立と幸福と驚きを 2 度ずつとした。これは、基礎的な表情の組み合わせとして、非慣習的な表情（6.6.6）や至極以外の表情は、被験者 N のために、認識処理をしていない。なお、被験者 N の顔の上部の有効フレーム数が 0 となっているが、これは、被験者 N が驚きの表情の際に全く眉を動かしていないためである。しかし、目と口によって驚きを表現しているので、図 9のように顔全体としては驚きの表情に見える。

表からわかるように、特微量抽出で94%、表情認識で90%弱と、高い成功率であった。特微量抽出時の失敗はすべて顔器官の抽出段階のものであり、特に眉や目領域の抽出に失敗しているものが多かった。眉や目は造作が細かく、かつ、頭髪などへの影響を受けやすいことが原因と考えられる。表情認識では、主に中立の表情から変化していく過程、あるいは中立の表情へ戻る過程での認識が多かった。特に、形態的な特徴が類似している驚きと恐怖を混同することが多かった。

入力画像の一例として、図7に被験者Nの「中立→幸福→中立」の表情変化を示す。また、この際の顔の下部について、フエジイ推論器からの表情認識出力例を図8に示す。図7からわかるように、5フレーム目前後から幸福の表情が始まり、45フレーム目前後まで継続してきている。この際の認識出力値は、5フレーム目前後から幸福の出力値が上がり、逆に中立が下がっている。45フレーム目前後は上がり、幸福の出力値が下がり、中立の出力値が上がりつつある。また、図中に示されているように、他の表情の出力値はほとんど変化していない。

表3に、被験者3名の表情認識実験の結果を示す。特徴量抽出に関しては、認識処理を行った全フレーム数のうち、各特徴量の抽出に成功したフレーム数を成功として示している。表情認識に關しては、幸福あるいは驚きの表情が検出されており、かつ特徴量の抽出に成功したフレーム数(有効)のうち、顔の上部、中部、下部について、その表情の認識能力が一番高かったフレーム数を成功として示している。幸福の表情では顔の上部に特徴的な動作で示している。

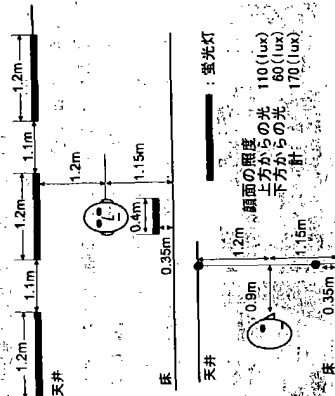


図6 評価実験時の照明環境

Fig. 6 Lighting Environment when Evaluation Experiment

30

表3 表情認識實驗結果

Table 3 Experimental Results of Facial Expression Recognition

友情	特種兵抽出 全數	拉拔者 A			
		上野	吉岡正隆		下野
		成功	有妨	成功	有妨
志保	31	—	—	220	24
幸助	39	—	—	22	24
真子	14	12	12	11	11
真子	1	11	12	23	11
真子	2	10	11	8	10
真子	2	25	11	6	10
計	104	112	23	61	32

去青		特種產出		臺灣特級綠					
				上青		中安綠		下青	
成	全	成	全	成	全	成	全	成	全
青綠	46	47	—	—	—	38	39	36	39
青藍	61	61	—	—	—	52	53	52	53
青白	27	35	5	6	5	6	7	7	8
青黃	26	33	7	9	7	9	8	10	11
計	160	177	12	15	102	107	103	110	—

表情	特種票抽出		表情發露				下野
	成功	全數	成功	上野	成功	中央部	
幸運1	58	58	28	38	37	40	
苦勞2	45	49	—	—	25	32	35
苦勞3	51	51	0	0	23	25	21
苦勞4	45	46	0	0	23	25	24
苦勞5	46	46	0	0	23	25	24
合計	199	204	0	0	101	121	117
							132

要情	金雞縣者					
	表報脫繳					
特種品抽出	上郡	中央部	下郡	上郡	中央部	下郡
成勢 / 全勢	280 / 289	185 / 210	189 / 207	成勢 / 有勢	成勢 / 有勢	成勢 / 有勢
華蘭計	183 / 204	34 / 38	79 / 88	83 / 96		
實金計	463 / 493	34 / 38	264 / 298	271 / 303		
成勢率	93.9 %	88.5 %	88.0 %	89.4 %		

表 4 1フレームの平均処理時間

Frame	處理	時間 (ms)
	顏領域抽出	13.3
	顏器官抽出	7.4
	特徵點抽出	6.1
	特徵量算出	-
	表情認識	5.0
	表情認識計	31.8

この評価実験では、特徴点抽出などの画像認識精度や心理速度については評価できているものの、被験者が男大学生3名だけであり、性別や年齢の違いによる表情認識精度については評価できていない。また、認識した表情も基本6表情のうち6つの幸福だけであり、他の表情については評価できていない。今後は、被験者数や表情を増やし、認識精度を再評価するとともに、問題点を明らかにして認識手法を改良していく必要がある。

5. 16 2 8

顔の表情は人間同士のコミュニケーションにおいて非常に重要な情報であり、人間は時間とともに変化している相手の表情から内面状態を読み取ることで、円滑

— 31 —



図9 被験者Nの「驚き」の表情

なコミュニケーションを実現している。人間と機械とのよりよいコミュニケーション方法を確立するためには、この人間の動的な表情変化を機械が認識し、人間の内部状態を推定する必要があると考えられる。そこで、本研究では、まずその基礎研究として、機械による動的表情認識手法を提案し、その手法に基づいてCCDカメラで撮影した顔面像からリアルタイムで表情を認識するシステムを試作し、評価実験により、その認識精度と処理速度を検討した。

本研究で提案した表情認識手法は、主にリアルタイム性を重視しており、処理速度の向上のため以下のような工夫を行った。

- AUをよく表現するだけでなく、画像処理により抽出が容易な部分を特徴点として用いている
 - 特徴量から表情を認識する際に、処理が簡単なフジ・ジ推論を用いている
 - 顔領域抽出、顔器官抽出の際に、縮小画像を用いている
- また、この表情認識手法は、1章で述べた他の表情認識のアプローチと比較して次のような長所がある。
- 画像認識処理方法が比較的単純であり、顔表情の変化をリアルタイムで認識処理できる
 - 画像処理の最初に顔領域を抽出しているため、背景や顔の動きに影響されにくい
 - 顔器官に設定した特徴点の位置変化を表情認識の手かりとしているため、年齢、性別、個人差の影響されにくい

● 顔にある程度均一に光が当たっている状態でない、顔器官の影などの影響により表情を正しく認識できない

- ファジィ推論器へ入力する特徴量の規格化のため、あらかじめ特徴量の最大値、最小値、中立の表情の際の標準値が必要である
- 本研究では、最終目標である表情の動的な変化から複雑な感情を推定するまでには至っていない。今後は、

国弘 威

影響環境に左右されないロボットの高い表情認識手法を開発し、さらに、本試作システムでは対象としていない視線方向、瞬き、頭部運動などでもあわせて認識するとともに、表情と感情との関連についてより広範な調査を行い、感情を推定できるシステムへと発展させたい。

謝 辞

本研究は、日本学術振興会、平成9年度未来型学術研究推進事業(JSPS-RFTF9700102)の一環として遂行したことを附記し感謝の意を表す。

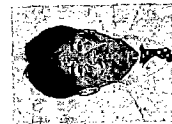
参 考 文 献

- [1] 松野 他: ポテンシャルネットとKL展開を用いた顔表情の認識, 信学誌(D-II), Vol. J77-D-II, No.8, pp.1591-1600 (1994)
- [2] 開瀬: オプティカルプロローブ抽出による表情筋の動作検出, 信学技法, PRU89-128, pp.17-24 (1990)
- [3] 中川, 金子, 原島: 低空間周波数成分を用いた人物顔表情の即時認識, 映像メディア処理シンポジウム(IMPS96), I-8.1, pp.57-58 (1996)
- [4] 大久保, 渡辺, 山田: 表情を用いたオプティカル・ネットワークによる口唇運動抽出, 第14回ヒューマン・インタフェース・シンポジウム論文集, pp.176-179 (1998)
- [5] ズデネク, 伊藤, 岡本: 動き情報を用いた表情特徴の自動抽出手法, 信学誌(D-II), Vol. J81-D-II, No.6, pp.1150-1159 (1998)
- [6] Ekman, P., Friesen, W.V.: The Facial Action Coding System, Consulting Psychologists Press (1978)
- [7] Ekman, P., Friesen, W.V.: 工訳: 表情分析入門, 誠信堂 (1987)
- [8] 小林, 原: ニューラルネットワークによる人の基本表情認識, 計測自動制御学会論文集, Vol.29, No.1, pp.112-118 (1993)
- [9] 佐々木, 赤松, 天永: 顔画像認識のための色情報を用いた顔の位置合わせ法, 信学技法, IE91-2, pp.9-15 (1991)

(1998年11月2日受付, 1999年3月18日再受付)

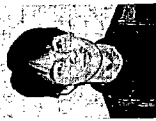
著 者 紹 介

下田 宏 (正会員)



1989年京都大学大学院工学研究科修士課程修了。同年(株)島津製作所に入社。主に医用画像診断装置の研究開発に従事。1996年京都大学大学院エレクトロニクス研究科助手、現在に至る。ヒューマンインタフェース、インタフェースデータベースに興味を持つ。日本原子力学会、システム制御情報学会会員。博士(工学)。

1997年京都大学工学部電気第二工学科卒業、現在、同大学院エレクトロニクス研究科修士課程在学。画像処理、顔表情によるコミュニケーションに興味を持つ。



吉川 榮和 (正会員)



1970年京都大学大学院工学研究科電気工学第二専攻修士課程修了。同年同大学原子エレクトロニクス研究科助手。1974年動力炉・核燃料開発事業団勤務。1981年京都大学原子エレクトロニクス研究科教授。1992年同教授。1996年同大学大学院エレクトロニクス研究科教授。現在に至る。エレクトロニクス情報学、マンマシンシステムの研究に従事。1992年システム制御情報学会基本記念賞論文賞、1995年計測自動制御学会論文賞受賞。電気学会、計測自動制御学会、日本原子力学会などの会員。工学博士。

ヒューマンインタフェース学会論文誌 投稿規程

1. 総 則

- (1) ヒューマンインタフェース学会論文誌(以下、論文誌と呼ぶ)の投稿はこの規程による。
- (2) 投稿原稿の著者のうち少なくとも1名は本学会会員とす。ただし、本学会論文誌編集委員会(以下、編集委員会)が承認または原稿執筆を依頼したものはこの限りではない。
- (3) 投稿原稿は、他の学術刊行物に公表されてはいないものに限り、学術刊行物とは、学協会または権威ある審査機関による審査を経た論文などが掲載されている刊行物で、通常の手段で入手が可能なものとする。
- (4) 論文誌に掲載された記事の著作権は本学会に帰属する。著作権の取扱詳細については、本規程の「著作権について」の「内規」に従うものとする。
- (5) 論文誌に掲載された記事内容についての責任は、すべて著者が負うものとする。

2. 投稿原稿

- (1) 原稿は和文または英文とする。
- (2) 記事の種類は、「原著論文」、「総説論文」、「技術報告」、「ジョイントノート」とする。このうち「ジョイントノート」はさらに「研究速報」と「紙上討論」を含む。
- (3) 記事の内容は、ヒューマンインタフェースに関連するものとする。
- (4) 原稿は、本「投稿規程」と「原稿作成の手引き」に従って執筆し、「投稿の手引き」に従って投稿するものとする。これらに準拠していない原稿は受理しないことがある。

3. 種 別

- (1) 原著論文(Original Paper)
(6ページを標準とする偶数ページ)
ヒューマンインタフェース分野における新しい研究・開発の成果の記述で、研究の対象・方法あるいは結果に独創性・創造性があり、かつ明確で価値のある結果や事実を含むもの。

- (2) 総説論文(Review Paper)
(6ページを標準とする偶数ページ)

対象とする分野における研究の、歴史的背景、重要性、進捗状況、今後の発展の方向などを踏まえつつ、著者の学術的、技術的研究あるいは開発成果を総合的にまとめたもの。

- (3) 技術報告(Technical Report)
(6ページを標準とする偶数ページ)

設備・装置、ソフトウェアなどの設計、制作、試験、運用、解析、評価などの新しい経験やその結果の報告で、実用的価値のあるもの。

- (4) ショートノート(Short Note)
(4ページ以内の偶数ページ)

(a) 研究速報 研究・技術上の新しい成果の速報、または技術上の新しい提案、後日、その詳細を原著論文として投稿することができる。

(b) 紙上討論 論文誌に掲載された記事に対する質問とそれに対する回答、討論で、会員にとって価値のあるもの。

4. 投 稿

- (1) 論文誌に投稿するためには、「投稿の手引き」の提出書類一覧表に記載の書類一式を、本投稿規程「6.連絡先」へ送付するものとする。
- (2) 論文誌に投稿するための書類一式が本学会に到着した日付をもって原稿の受理日とする。
- (3) 投稿原稿の採否は、査読の結果に基づいて編集委員会が決定し、著者に通知する。編集委員会は投稿原稿について修正を求めることがある。
- (4) 査読照会事項に基づいて原稿の修正を行う場合は、旧原稿と査読所見に対する回答書を添えて、3ヶ月以内に書類一式を再提出する。旧原稿の返却後、期限内に再提出されない場合は期限切れにより原稿の撤回とみなされる。著者の都合により撤回する場合には、その旨を編集委員会まで書面で連絡する。撤回された原稿が再度投稿された場合は、新投稿扱いとする。
- (5) 投稿原稿の再提出において、原稿が大幅に修正された場合には、編集委員会の判断により新投稿扱いとすることがある。
- (6) 編集委員会において掲載が決定した後でも、やむを得ない事情により決定を取り消すことがある。この場合は、編集委員会がその理由をすみやかに

ヒューマンインタフェース学会誌 ISSN 1344-7254
ヒューマンインタフェース学会論文誌 ISSN 1344-7262

Human
Interface

- Optimum Apportionment of Alphabet-presentation Time with Time Lag between Alphabet Letters Sequentially Presented in a Horizontal Row and Alphabet Sound
YOKOI Hirokazu
- Study on Modeling Plant Operator's Anomaly Diagnosis and Deduction of Human Cognitive Reliability (HCR) Curve by Human Model Simulation
WU Wei, NAKAGAWA Takashi, YOSHIKAWA Hidekazu
- A Prototype of a Real-time Expression Recognition System from Dynamic Facial Image
SHIMODA Hiroshi, KUNIHIRO Takeshi, YOSHIKAWA Hidekazu